

The Future of Online Expression and Innovation Depends on Robust Section 230 Protections

BY DAVID INSERRA

EXECUTIVE SUMMARY

Section 230 is a cornerstone of how online speech works in the United States. It allows websites and platforms to host user-generated content without being legally responsible for everything their users say, while still giving them the freedom to remove content that they and their users may find objectionable. This legal framework helped the internet grow into a place where millions of Americans can speak, create, organize, and do business.

- Section 230 removed legal incentives that discouraged content moderation, allowing online platforms to engage in the moderation of user-generated content without assuming liability for all user speech.

- Section 230 has enabled *more* free expression and innovation by protecting platforms from being held legally responsible for everything posted by their users.
- Weakening or repealing Section 230 would backfire, pushing platforms to over-remove lawful speech, limit features, or exit markets to manage legal risk.
- Expanding platform liability would entrench large incumbents by raising legal and compliance costs that small and new companies cannot absorb, ultimately reducing competition and innovation.
- As new technologies emerge, such as generative AI, preserving a stable speech-protective liability framework is essential for the United States to remain a leader in innovation and free expression.



DAVID INSERRA is a fellow for free expression and technology at the Cato Institute.

INTRODUCTION

The US has the world’s most vibrant online speech industry. This industry has arisen from the incredible innovation of new products and platforms and resulted in a drastic increase in the ability of users to create content, express themselves, and access information. The explosion in innovation and expression are the result of the US’s pro-innovation policies.

Perhaps the most important and oft-discussed policy is Section 230 of the Communications Act, as amended by the Telecommunications Act of 1996. Section 230 provides liability protection for online platforms by holding those who post content online responsible for any harms associated with online speech, rather than the platforms themselves. In the 30 years since the passage of Section 230, countless websites, tools, and platforms have emerged and prospered as a result.

But not everyone has seen Section 230 as an undisputed good.¹ Some have framed Section 230’s liability protections as a special handout to big technology companies.² Others blame section 230 for the spread of arguably harmful content online.³ Still others believe it is used to bias online speech against one group or political party.⁴ And many have confused how Section 230 interacts with the First Amendment’s protection of free expression, even 30 years after Section 230 went into effect.⁵ For these reasons, there have been many proposals to modify or eliminate Section 230.

“While policymakers and critics may have various concerns around technology, Section 230 is not the problem, but rather part of the solution.”

In this paper I will explain the history behind Section 230, what it does, and how it works together with the First Amendment to protect speech online. I will also show why various efforts to limit or remove Section 230 would be harmful to expression and innovation, and discuss how we should think about Section 230 as we move forward with new technologies, such as AI.

While policymakers and critics may have various concerns around technology, Section 230 is not the problem, but rather part of the solution. By protecting online platforms—big and especially the small—from onerous legal liability, Section 230 continues to move us toward a more innovative and free future.

A BRIEF HISTORY OF ONLINE SPEECH AND LEGAL PROTECTION

The early 1990s saw the rise of the modern internet with the development of the World Wide Web and the emergence of user-friendly browsers.⁶ But with the emergence of this new boundless internet, legal questions quickly arose regarding intermediary liability, that is, to what degree are intermediaries such as websites liable for hosting defamatory and other forms of problematic content created by third parties.

Why Section 230 Was Created

In the case of *Cubby, Inc. v. CompuServe Inc.* (1991), Cubby was an early online news publication that sued a competitor newsletter, Rumorville, for libelous statements made about Cubby’s business practices. Cubby also sued CompuServe—one of the dominant early online services that connected users online via message boards, moderated forums, file sharing, and connection to the broader internet—because CompuServe hosted the Rumorville publication.⁷ Ultimately, the federal court dismissed the case since CompuServe was a distributor and “neither knew nor had reason to know of the allegedly defamatory statements.”⁸ Given the growing amount of content on the nascent internet, the decision recognized that it would be “no more feasible for CompuServe [or other internet services and platforms] to examine every publication it carries for potentially defamatory statements than it would be for any other distributor to do so.”⁹

But even this level of liability protection was only partial. Once a distributor was notified of the potentially problematic speech, the distributor then had reason to know and could be held liable for any inaction from that point forward.¹⁰ In essence, the *Cubby* case had created a notice and takedown liability regime. And it created an incentive

for online services to be ignorant about the content they were hosting because then they could legitimately claim to have no knowledge of any given piece of content and thus avoid liability.

“Companies may face a ‘moderator’s dilemma,’ in which they have a perverse incentive to not moderate, even if most users want at least some moderation.”

And then the *Stratton Oakmont, Inc. v. Prodigy Services Co.* case in 1995 would further throw the question of legal liability into doubt. A user posted on Prodigy’s “Money Talk” message board accusing Stratton Oakmont and its president, Daniel Porush, of various financial crimes and fraud—notable given that Porush would plead guilty to financial crimes in 1999, and Stratton Oakmont is the company depicted in film *The Wolf of Wall Street*. But Porush sued Prodigy, arguing that Prodigy should be held liable for defamatory content posted on its message boards. The primary difference between Prodigy and CompuServe was that Prodigy tried to play some role in moderating content. It had content guidelines and could remove posts or users for violations.¹¹ The New York court found that “such decisions constitute editorial control” and that even though such control was incomplete or imperfect, it “does not minimize or eviscerate the simple fact that Prodigy has uniquely arrogated to itself the role of determining what is proper” on its service.¹² By trying to be a more “family-oriented” type of service, Prodigy had made itself liable for content posted by its users.¹³

These cases presented a dilemma to the burgeoning set of websites hosting user speech. On the one hand, platforms may want to remove obscene, noxious, or objectionable content to cultivate websites that do not scare away or disgust users, but engaging in such moderation makes them liable for any harmful speech. On the other hand, being purposefully ignorant of user-generated content and moving to aggressively remove content only once informed of its potential harm might create disruptive and unhelpful experiences for users, but at least companies acting this way

couldn’t be held liable for what their users posted. The result of this “moderator’s dilemma” was that companies faced a perverse incentive to not moderate, even if most users wanted at least some moderation.¹⁴

In response to the moderator’s dilemma exemplified by the *Cubby* and *Stratton Oakmont* decisions and concerns about the impact they would have on online speech and innovation, Reps. Chris Cox (R-CA) and Ron Wyden (D-OR) successfully added an amendment to the Communications Decency Act of 1996, a law attempting to regulate indecent online speech through updates to the Communications Act of 1934.¹⁵ This amendment would become Section 230 of the Communications Act.

What Is Section 230?

Section 230 of the Communications Act opens with the findings of Congress that identify the importance of allowing the internet to flourish through various online services and the desire of the US government to promote innovation and greater expression by removing disincentives for the development of content-moderation tools. Even years later, this remains central to understanding the law. The core legal protections include:

- (c) Protection for “Good Samaritan” Blocking and Screening of Offensive Material.—
 - (1) Treatment of Publisher or Speaker.—No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.
 - (2) Civil Liability.—No provider or user of an interactive computer service shall be held liable on account of—
 - (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
 - (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in [subparagraph (A)].¹⁶

What Liability Protections Section 230 Offers

Section 230 offers two related sets of liability protection to all websites, platforms, and online service providers that host user-generated content.¹⁷ First, Section 230 states that interactive computer services shall not be treated as the publisher or speakers of information that was provided by a third party. This provision is aimed at the limited-liability protection offered under the *Cubby* decision by clearly establishing that user-generated content that an online service hosts, publishes, or distributes should not be treated as its own speech. As Cox has summarized: “What Section 230 added to the general body of law was the principle that an individual or entity operating a website should not, in addition to its own legal responsibilities, be required to monitor all of the content created by third parties and thereby become derivatively liable for the illegal acts of others.”¹⁸

“The innovation of Section 230—notes Rep. Chris Cox (R-CA)—is that a website should not be required to monitor all of the content created by third parties and become derivatively liable for the illegal acts of others.”

The second liability protection offered by Section 230 is that a platform maintains its protections from liability even if it acts to moderate content. This provision directly addresses and reverses the *Stratton Oakmont* decision by allowing platforms to moderate online content in good faith that the platform considers obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, without being held liable as a result. This provision also protects platforms that create and share tools that empower others to restrict content that may be objectionable. In sum, it permits companies to create websites and platforms that are not a free-fire zone where any and all speech must be tolerated. Instead, it allows companies to take action against speech that is protected by the First Amendment but that the company and its users

may not wish to see. There has been some disagreement among the relatively few court cases addressing how widely to interpret the “good faith” or “otherwise objectionable” clauses, but the decisions of these cases have generally supported the right to reduce or remove content.¹⁹

Furthermore, Section 230 exempts several categories from liability protection. Section 230 cannot be used as a defense in cases involving the enforcement of federal criminal laws, which has been interpreted to mean that Section 230 provides protection from prosecution under state criminal laws, but not federal law.²⁰ For example, Section 230 does not shield platforms from prosecution under child sexual exploitation laws if the federal government finds a platform to be actively taking part in the exploitation of children. Section 230 immunity also does not apply to “any law pertaining to intellectual property,” thus it does not protect against lawsuits regarding copyright or trademark issues.²¹

Essential Legal Interpretations of Section 230

The first legal case involving Section 230 was *Reno v. ACLU*, in 1997, in which a unanimous Supreme Court quickly struck down most of the Communications Decency Act for violating the First Amendment. But importantly, it did not strike down Section 230.²² As a result, Section 230 went into effect and began to appear in legal disputes.

The first major case focused on Section 230’s protections was *Zeran v. America Online, Inc.*, also in 1997. In 1995, an anonymous user or users posted content on America Online’s (AOL’s) bulletin boards that praised the Oklahoma City Bombings, notably by advertising for sale items that included slogans such as “Visit Oklahoma, It’s a BLAST!!!” or “McVeigh for President 1996.” The advertisements told those who were interested in the items to call Ken, and they listed the phone number of Kenneth Zeran, who had nothing to do with the content. Zeran was quickly inundated with phone calls expressing anger and threats. While AOL worked to take down content when notified by Zeran, it could not stem the tide of new posts that resulted in more phone calls and threats. Zeran sued AOL for not taking sufficient action to remove this harmful content. But the district court and the Fourth Circuit agreed that Zeran was attempting to hold AOL responsible for being a publisher of content that

it had not created. The courts in *Zeran* faithfully interpreted Section 230, stating that

Congress made a policy choice, however, not to deter harmful online speech through the separate route of imposing tort liability on companies that serve as intermediaries for other parties' potentially injurious messages. Congress' purpose in providing the § 230 immunity was thus evident. Interactive computer services have millions of users. . . . The specter of tort liability in an area of such prolific speech would have an obvious chilling effect. . . . Faced with potential liability for each message republished by their services, interactive computer service providers might choose to severely restrict the number and type of messages posted.²³

In *Zeran* and subsequent cases, the courts' Section 230 jurisprudence recognized that liability protections allow platforms to have lawsuits that target them for hosting and moderating the speech of third parties dismissed quickly, rather than going through the full range of costly litigation. This interpretation was formalized by the courts into a three-part test, first developed in *Barnes v. Yahoo!*, when considering whether to dismiss a lawsuit on the grounds of Section 230:

- The defendant is an interactive computer service;
- The plaintiff's claim treats the defendant as a publisher or speaker; and
- The plaintiff's claim derives from content the defendant did not create.²⁴

If a platform meets all these conditions, then it is protected from liability and the case is dismissed.

Section 230's Benefits to Companies

Without Section 230, some lawsuits would find companies liable for hosting different sorts of harmful speech or making certain content-moderation decisions, thus imposing significant penalties on them. Other lawsuits against platforms would ultimately fail to prove that the platform was responsible for whatever harm was alleged, yet would still be very costly.

According to a 2019 report by Engine, an organization focused on supporting start-up technology companies, legal costs dramatically increase once litigation begins. The costs of discovery and an actual court battle can rise into hundreds of thousands of dollars or more.²⁵ Even for an organization confident of winning at trial, it may simply be cheaper to settle rather than proceed. The problem is that facing such legal costs as a regular occurrence will quickly devastate a company. Section 230's liability protections allow platforms to request the dismissal of such cases right from the beginning. The legal costs of handling a motion to dismiss are likely in the \$15,000 to \$40,000 range but can reach as high as \$80,000. Still, such costs are far lower than a full lawsuit and, because it is handled early in the legal process, succeeding with a motion to dismiss will be far less disruptive to the ongoing operations of the platform. And Section 230's clear protections also serve as a disincentive for lawsuits targeting platforms for hosting user-generated content in the first place.

“Legal protections are especially critical for smaller platforms and websites that simply do not have the resources to employ extensive content moderation and legal teams.”

These legal protections are especially critical for smaller platforms and websites that simply do not have the resources to employ extensive content moderation and legal teams that would be necessary to adjudicate cases in a more litigious environment without Section 230.²⁶ Without Section 230, these small companies would have to pull their limited staff and monetary resources away from developing innovative new products and services to instead address constant legal risks and costs. This may entirely prevent a startup from developing a successful product. A 2019 study by Techdirt founder Mike Masnick found that 28 percent of social media startups—governed by Section 230—were successful. But startups in the digital music space—governed by a far less protective liability regime under the Digital Millennium Copyright

Act—succeeded a bit less than 8 percent of the time.²⁷ Section 230 was, and remains, essential for online innovation and competition.

CHALLENGES TO SECTION 230 AND ONLINE SPEECH

Section 230, together with the First Amendment’s protections for speech, has created a powerful bulwark against government and costly legal threats to online speech, resulting in an unprecedented surge in expression and economic opportunity. Such changes, however, have given rise to various objections, whether it be complaints about too little content moderation, too much or biased moderation, or just that Section 230 is a special handout to technology companies that is unfair and unnecessary.

Ultimately, even though these issues vary significantly, they have created enough common cause for a significant number of experts and policymakers to call for major reform, or even abolition, of Section 230, as well as more general assaults on online speech. Over time, these critics have proposed a host of potential changes or challenges to Section 230 or the First Amendment, with several limited examples passing into law. While they may be well intentioned, such changes pose a major risk to free expression and innovation by threatening intermediaries that carry user-generated content. These various arguments and proposals may seek to rectify some real or perceived inequity or injustice under the current pro-expression legal framework, but these proposals to create additional liability and responsibilities for online platforms will have significant counterproductive consequences.

The primary objections to Section 230 protection for online speech could be grouped into three buckets: failures to moderate; aggressive moderation and bias; and Section 230 functioning as a special handout to large companies

Failures to Moderate

A core critique of Section 230 is that platforms have failed to moderate sufficiently to merit legal protection. This criticism sometimes targets specific categories of content—such as pornography, human trafficking, or terrorist material—and at other times it reflects broader

concerns, including content deemed harmful to minors. While free-expression advocates rightly emphasize the extraordinary benefits that strong legal protections have enabled, these arguments do not negate the seriousness of these concerns.

One of the most prominent challenges to Section 230 involved Backpage.com, a classified advertising site similar to Craigslist, which became known for its adult services section. That section frequently included sex trafficking, including trafficking of minors. Although Backpage claimed to combat illegal activity, it often edited ads to obscure their illegality, effectively facilitating trafficking. While courts initially shielded Backpage under Section 230—sometimes expressing regret at their inability to impose accountability—later rulings found that the site’s active role in rewriting ads made it an information provider, thus stripping it of Section 230 protection. The site’s owners were ultimately convicted of money laundering under existing law in 2023.²⁸

“The most immediate and predictable consequence of measures intended to hold companies liable for harmful content is to strongly incentivize overmoderation in order to reduce legal risk.”

In two cases that were decided by the Supreme Court in 2023, *Google v. Gonzalez* and *Twitter v. Taamneh*, families of victims killed in ISIS terrorist attacks argued that social media platforms aided and abetted terrorism by allowing terrorist content to spread. The platforms contended that Section 230 barred such claims and that merely operating platforms that were misused by terrorists did not constitute aiding and abetting. The Court declined to rule on Section 230, instead holding that a failure to remove enough ISIS-related content—“out of hundreds of millions of users worldwide and an immense ocean of content”—did not amount to intentional aid or participation in terrorism.²⁹

The situations described here and elsewhere are unquestionably tragic.³⁰ In response, policymakers have proposed—and in some cases adopted—measures intended

to hold companies accountable for harmful content. The most immediate and predictable consequence of these policies, however, is to strongly incentivize over-moderation in order to reduce legal risk. Platforms must identify and remove content for which they may be held liable. Yet beyond such clearly covered content lies a far larger gray area, where users discuss harmful topics in indirect, academic, satirical, critical, or otherwise lawful or nonharmful ways. Although this speech is not the intended target of such policies, its proximity to prohibited content makes it risky to host. Faced with limited moderation resources and significant legal exposure, companies will predictably err on the side of removal, silencing large amounts of legitimate and valuable speech at scale.

The Error of FOSTA–SESTA

A prominent example of over-moderation following a policy change that limited Section 230’s liability protections is the 2018 joint passage of the Fight Online Sex Trafficking Act and the Stop Enabling Sex Traffickers Act (FOSTA–SESTA, commonly referred to as just FOSTA). FOSTA was enacted largely in response to Backpage, even though courts ultimately allowed Backpage to be prosecuted and taken down under existing federal criminal law. Early prosecutorial failures, however, led policymakers to conclude that additional action was needed to compel platforms to do more to combat sex trafficking. No one doubts the good intentions behind these efforts to stop horrific abuse.

“As with many online speech issues, existing laws are often sufficient to address harmful conduct, although common-law development may take time to reach the correct outcome.”

But sadly, the actual impact of FOSTA has been to broadly silence online conversations about sex.³¹ By stripping Section 230 protections from platforms hosting content related to sex trafficking and prostitution, the law exposed them to significant legal risk. As a result, platforms have

removed lawful and nonharmful content to avoid liability. This includes sex education, discussions among sex workers about safety or dangerous clients, and information about legal sexual activity. In some cases, entire sex-related platforms shut down in response to liability concerns. Rather than protecting people from sexual crimes or trafficking, FOSTA has often limited access to safety-related information. Moreover, FOSTA has rarely been used to prosecute traffickers and it has not been successfully used to impose civil liability on platforms.³²

Finally, this rarely used law—with its narrow but outsized impact on lawful speech—was not necessary to address Backpage. Courts ultimately determined that Backpage’s conduct fell outside Section 230’s protections and convicted its owners of money laundering related to prostitution and sex trafficking. As with many online speech issues, existing laws are often sufficient to address harmful conduct, although common-law development may take time to reach the correct outcome.

Broader Proposals to Compel Moderation

Other proposed bills would similarly create liability for platforms that host various forms of user-generated content, including platforms that:

- host the advertisement and sale of illicit drugs;³³
- host firearm sales;³⁴
- fail to report terrorist activity;³⁵
- host “health misinformation;”³⁶
- fail to remove any content that could cause “irreparable harm;”³⁷
- use algorithms to recommend or sort content that is harmful;³⁸
- host content resulting in “bodily injury or harm to mental health” of children;³⁹ or
- fail to take “reasonable efforts to address unlawful activity.”⁴⁰

Of course, some advocates and policymakers have proposed going as far as removing Section 230 entirely.⁴¹

Each of these proposals would create new areas of potential liability for platforms and websites that extend beyond criminal and objectively harmful content, especially

as some harms being targeted are vague and subjective. For example, removing liability protections for platforms hosting content related to illicit gun sales or gun crimes would likely result in the removal of lawful discussion of firearms or gun rights advocacy. Eliminating Section 230 protections for platforms where illicit drug sales occur would similarly incentivize the removal of content about drug rehabilitation and recovery. Policies aimed at protecting minors from harmful content, such as material related to anorexia or self-harm, would likely suppress access to resources and support for those facing mental health challenges. In each case, the threat of liability may reduce some harmful content, but it will also predictably eliminate lawful and even beneficial speech, sometimes closing entire communities, websites, or platforms.

“The threat of liability may reduce some harmful content, but it will also predictably eliminate lawful and even beneficial speech, sometimes closing entire communities, websites, or platforms.”

The effects would extend far beyond social media, shaping what products can be sold on platforms such as Amazon or Etsy, what reviews appear on Yelp or TripAdvisor, and what search results are displayed by Google or DuckDuckGo. Proposals to repeal Section 230 altogether would fundamentally alter online engagement, reviving the perverse incentives of the early 1990s: Platforms would either moderate very little or else aggressively moderate, quickly removing any content that poses even a potential liability.

A recent example from 2025 is the well-intentioned TAKE IT DOWN Act, which holds platforms responsible for failing to promptly remove nonconsensual intimate imagery, including AI-generated content, despite Section 230’s traditional protections for hosting third-party content. While such material is illegal and those who create it should be held accountable, imposing liability on platforms incentivizes more moderation than likely would otherwise occur. Under a

notice-and-takedown regime, platforms will likely remove a wide range of sexual content out of fear of Federal Trade Commission enforcement, including political satire, cartoon or AI-generated imagery, controversial but lawful material, or sexual content that someone simply wants suppressed. Although such speech may fall outside the law’s scope, platforms would still have to assess legal risk within 48 hours.

Compounding the problem, the reporting system imposes no penalties for fraudulent or frivolous claims, thus enabling abuse by political or powerful actors. Smaller platforms, with limited moderation capacity and legal resources, will struggle the most and may respond by removing speech by default or banning entire categories of content to reduce their risk.

Section 230 Critiques Are Often Attacks on the First Amendment

It is important to also note that many policies target third-party speech that, while potentially harmful or illegal, includes a great deal of speech protected by the First Amendment. Content labeled as “misinformation,” generically harmful content, and firearm sales are examples of protected speech. In these instances, policymakers aren’t really frustrated with Section 230—they are frustrated with the First Amendment.

This frustration was recently on display in the aftermath of the assassination of Charlie Kirk, in which some users posted content ghoulishly celebrating the murder. This led Rep. Clay Higgins (R-LA) to demand that social media companies remove such content and ban users from posting it. Higgins argued: “If you shield these offenders, Section 230 will not protect your platform from vigorous exposure.”⁴² Higgins’ anger towards Section 230 is misplaced; while platforms can remove this content if they wish to, the speech he objects to is protected by the First Amendment, and it would similarly be lawful if it were written in a newspaper or stated in a city council meeting.⁴³

Platforms Are Still Responsible for Their Own Speech

Section 230 does not protect platforms, however, from liability for everything they do. As discussed earlier, courts ultimately found the owners of Backpage guilty for their role

in sex trafficking because their actions materially contributed to the development of unlawful content. Section 230 defines an “information content provider”—and thus a potentially liable party—as any entity “responsible, in whole or in part, for the creation or development of information.”⁴⁴ While early cases such as *Zeran* established the statute’s core interpretation, courts have since worked to clarify when a platform’s conduct transforms it from a protected interactive computer service into an unprotected content provider.

One of the first major cases to address this question was *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC* (2008).⁴⁵ There, the Ninth Circuit held that Roommates was potentially liable for discriminatory content because it created mandatory questions and pre-filled dropdown answers that enabled illegal discrimination. Because those questions and answers were the platform’s own speech, Section 230 did not apply. But the case noted that Roommates did maintain Section 230 liability protection for other parts of its form that did not include Roommates’ speech, such as user-generated content entered into the “additional comments” field. This distinction attempted to put into practice Section 230’s key principle: Platforms are not responsible for carrying or moderating user speech, but they are responsible for their own.

Subsequent cases reflect continued disagreement over which platform activities transform companies into potentially liable speakers. A prominent example is *Herrick v. Grindr* (2019). Matthew Herrick’s former boyfriend created false dating profiles impersonating Herrick, leading more than 1,400 men to appear at Herrick’s home and workplace in the belief that he was seeking sexual encounters. While several platforms removed the impersonating accounts, Grindr did not. Herrick sued, arguing that Grindr’s product and design decisions were unsafe, particularly its failure to implement tools to block virtual private networks (VPNs) and prevent the impersonation. Grindr responded that its design choices—including allowing VPN access—were core content moderation and curation decisions protected by Section 230. Given that homosexuality is illegal or stigmatized in many jurisdictions around the world, Grindr deliberately allowed VPN access to protect user privacy and security. The Second Circuit agreed, holding that Herrick’s injuries were caused by his ex-boyfriend’s actions and that the lawsuit impermissibly sought to hold Grindr liable for user-generated content.⁴⁶

By contrast, in *Lemmon v. Snap* (2021), the Ninth Circuit held that Section 230 did not apply where Snapchat created a speed filter that appeared to reward fast driving and was used in a fatal high-speed crash. The court concluded that Snapchat was subject to negligent design claims because it created the filter, and that the plaintiffs’ claims “standing independently of the content that Snapchat’s users create with the Speed Filter.”⁴⁷ That decision has divided Section 230 supporters. Some worry that *Lemmon* opens the door to lawsuits targeting core platform features, such as algorithms, recommendations, groups, or content categorization. As legal scholar Eric Goldman notes, the case leaves uncertainty about whether Section 230 would protect YouTube if it created an “urbexing” category that allegedly encouraged trespassing, or would protect TikTok if a user injured themselves or a bystander while performing a dangerous stunt that the platform was accused of promoting, even if the video was never posted.⁴⁸

“In many cases, third-party speech is protected by the First Amendment; in these instances, policymakers aren’t really frustrated with Section 230—they are frustrated with the First Amendment.”

The decision in *Lemmon* may leave the door open to broader product liability challenges to platforms’ basic designs; perhaps the distinction worth noting is not one of product liability that those like Goldman rightly fear, but rather that of speech creation and development as discussed in the *Roommates* and *Backpage* decisions. Snapchat created the speed filter, and the filter significantly modified and developed the final piece of content. Section 230 does not apply to any information content provider “that is responsible, in whole or in part, for the creation or development of information.”⁴⁹ So in this sense, the speed filter was Snapchat participating, at least in part, in the creation of information or content rather than merely hosting it or recommending it. Other product

design features that merely categorize, curate, recommend, and host content are different in that they share user content but don't participate in the creation of the content. And just because a company such as Snapchat does not have Section 230 protections for speed filters does not make it immediately liable, but courts can decide the degree to which speech-modifying tools, such as a speed filter, are responsible for some harm.

Taken together, these cases show courts engaging in a common-law process to define the boundaries of Section 230. The statute shields platforms from liability for hosting user-generated speech but not for speech they help create. Proposals to further limit or repeal Section 230 risk disrupting this balance, encouraging over-moderation, silencing lawful speech, and sometimes harming the very groups such measures aim to protect.

Accusations of Ideological Bias

Another common critique of Section 230 is that it has allowed large technology firms to consistently suppress or censor too much speech or certain viewpoints. This is an argument most frequently, but not exclusively, made by conservative and right-of-center figures.⁵⁰ Most notably, this was used to justify the passage of bills in Texas and Florida forbidding biased content moderation. Gov. Greg Abbott (R-TX) defended the legislation: "There is a dangerous movement by social media companies to silence conservative viewpoints and ideas. That is wrong, and we will not allow it in Texas."⁵¹ Sen. Ted Cruz (R-TX) noted: "There are a great many Americans who I think are deeply concerned that Facebook and other tech companies are engaged in a pervasive pattern of bias and political censorship."⁵² President Donald Trump, posting on X, said that "Republicans feel that Social Media Platforms totally silence conservatives voices. We will strongly regulate, or close them down, before we can ever allow this to happen."⁵³

These critiques argue that social media should be neutral and open to all perspectives. Online platforms, in this view, are not private venues but rather public spaces, like a town square, that should not be closed off. Some people making this argument repeat the mistaken understanding that under Section 230, online service providers that want

liability protection as platforms cannot also exercise their First Amendment rights to editorial control through content moderation and curation. As a result, the policy proposals coming from these critics try to control and limit content moderation either by placing conditions on Section 230's liability protection or by ignoring Section 230 and arguing that requirements on how social media moderate content do not run afoul of the First Amendment.

Policies to Make Platforms Carry Speech

There have been multiple proposals to directly limit Section 230 in order to discourage content moderation, either broadly or in specific context. These include bills to limit Section 230 moderation protections for moderation of illegal content and select other categories, such as terrorism promotion content, self-harm content, excessively violent content, or obscene content;⁵⁴ limit Section 230 protections only to the moderation of illegal content;⁵⁵ label social media platforms as common carriers, creating new obligations for platforms to carry user content, not allowing for discrimination, and creating private rights of action and state government rights to sue as well;⁵⁶ and fully repeal Section 230.⁵⁷

“The result of turning moderation decisions into a potential liability will be that platforms are forced to provide what is essentially an unfiltered and ugly social media experience to users.”

Such efforts to repeal or limit Section 230 would predictably recreate the perverse incentives that followed *Cubby*. Any moderation would carry potential liability, encouraging platforms to adopt a hands-off approach. Comment sections would be overwhelmed with vile content; social media flooded with spam, pornography, and graphic imagery; and niche platforms, such as Christian Mingle or Grindr, would be forced to host content their users find bigoted or immoral. Review sites such as TripAdvisor or Yelp would be compelled to

allow off-topic or abusive comments that undermine meaningful reviews. At the same time, those most eager to sue platforms would exploit these rules to suppress critical content. The result would be widespread closure or restriction of comment sections, the removal of useful reviews, distorted information for users, and entire websites shutting down due to compliance and legal costs.

As previously mentioned, Texas and Florida passed laws that prohibited large social media platforms from moderating content in certain ways. Florida required platforms to “apply censorship, deplatforming, and shadow banning standards in a consistent manner among its users on the platform.”⁵⁸ Florida also restricted the moderation of journalistic enterprises and candidates for office. Texas prohibited moderation based on “the viewpoint of the user or another person” or the viewpoint represented in the user’s expression.⁵⁹ Both Texas and Florida claimed to have exceptions that allowed for the moderation of exploitative, violent, and illegal content.

“By creating liability for moderating too much or in ways that a government believes to be censorial and biased, platforms face perverse incentives to relax all moderation.”

These laws ultimately were struck down by the Supreme Court, but the case revolved around the First Amendment and online speech—issues that the case did not fully resolve.⁶⁰ But in her majority opinion, Justice Kagan wrote that “the editorial judgments influencing the content of those [platform] feeds are, contrary to the Fifth Circuit’s view, protected expressive activity. And Texas may not interfere with those judgments simply because it would prefer a different mix of messages.”⁶¹

And just as with the federal legislation discussed above, by creating liability for moderating too much or in ways that a government believes to be censorial and biased, platforms face perverse incentives to relax all moderation. For example, the Texas law prohibits “censorship,” that is, content moderation, based on viewpoint. But the

Ku Klux Klan, neo-Nazis, Antifa, ISIS, and many other violent extremist organizations have viewpoints. If proponents of these viewpoints aren’t directly inciting violence but instead merely sharing their offensive ideas, the Texas law would prohibit platforms from moderating such content. The same risk applies to pro-anorexia, pro-suicide, and other harmful viewpoints. Moderating graphic content in a way that restricts pro- or anti-abortion content, human rights abuses, or transgender medical information could also trigger liability. And since these laws allow private actors to pursue lawsuits, nearly any perceived inequity in treatment could result in litigation.

The result again is that many moderation decisions will create a potential liability and platforms will be forced to provide what is essentially an unfiltered and ugly social media experience to users in those states.⁶² Alternatively, some platforms may be able to survive by moderating very strictly and disallowing entire categories of speech.⁶³ On platforms taking this approach, cat videos might survive but discussions of major social and political issues would be highly curtailed.

Other Attempts to Compel Platforms to Host Speech

On the regulatory front, there are also proposals to grant the Federal Communications Commission (FCC) authority to police the application of Section 230 by enforcing its good-faith moderation clause.⁶⁴ In practice, this easily could be used to punish insufficient moderation or decisions that conflict with the preferences of the current FCC. Because the Commission’s leadership changes with each administration, moderation deemed good faith today could be considered bad faith tomorrow, making the standard impossible to apply consistently.

Such a regime would likely produce sharp swings in content moderation every four years and normalize government jawboning, as informal pressure could be backed by the threat of formal FCC investigations and penalties.⁶⁵ This approach would also raise serious First Amendment concerns due to its inherently viewpoint-based enforcement, and it would also require an expansive reading of the FCC’s authority over Section 230—one that courts are unlikely to accept under post-*Chevron* jurisprudence.⁶⁶

As with those who want more moderation, those who want less moderation or less bias have also argued for policies that try to sidestep Section 230. For example, they might argue that rather than directly limiting Section 230, platforms have to promise that their platforms are open for discussion in their terms of service or marketing. The platforms, therefore, should be held responsible when they moderate aggressively or moderate in a way that is deemed to be biased, because they are ostensibly failing to live up to their end of the contract they made with users. Interestingly, this same argument is made by advocates who believe that platforms have failed to remove enough harmful speech.⁶⁷

Such arguments may cite *Barnes v. Yahoo*, in which a court found that Section 230 did not apply to a decision made by Yahoo. In that case, Yahoo had explicitly promised Barnes that they would remove some defamatory content. And Barnes, relying on that promise, then didn't take other actions to protect herself because she was relying on that promise. So when Yahoo broke its promise, Barnes suffered additional harms since she had relied on Yahoo's promise. In other words, the case did not hinge on broad terms of service obligation, but rather the legal doctrine of promissory estoppel.

Furthermore, looking at the nature of terms of service agreements quickly makes it clear that they rarely place strict requirements on how exactly social media platforms should moderate. For example, here are the relevant sections of Meta's Terms of Service regarding users expression and moderation:

We help you find and connect with people, groups, businesses, organizations, and others that matter to you across the Meta Products you use. . . .

There are many ways to express yourself on Facebook to communicate with friends, family, and others about what matters to you. . .

If we learn of [violating] content or conduct like this, we may take appropriate action based on our assessment that may include - notifying you, offering help, removing content, removing or restricting access to certain features, disabling an account, or contacting law enforcement. . . .

You may not use our Products to do or share anything:

That violates these Terms, the Community Standards, or other terms and policies that apply to your use of our Products.⁶⁸

The language regarding what Meta promises to do is quite nuanced and aspirational, rather than strict and binding. It says it will help users connect with others and express themselves. It says it may take down harmful content, but it depends on their policies and assessments of any individual piece of content, not to mention the countless numbers of mistakes that will naturally occur, as platforms are making billions of content-moderation decisions. On the other hand, it places a binding obligation on users to follow the content policies and other policies that Meta establishes as terms for using its platforms.⁶⁹ In other words, if a user is unhappy that Meta has moderated their content because of a violation, Meta has not broken its terms of service. Instead, it is the users that have broken the terms of service, either deliberately or because they have misunderstood Meta's policies.

Section 230 as a Special Handout

In addition to arguments about whether existing laws protecting content moderation are incentivizing poor or biased behavior, a final line of argumentation is that Section 230 somehow is no longer necessary or is a special handout to large technology companies.⁷⁰ In other cases, these arguments seem to suggest that what was perhaps needed at the start of an infant industry is now problematic.

These arguments are usually used as additional justification for why platforms should lose Section 230 protection. Ultimately these arguments purport that Section 230 protections have, at best, distorted online speech and content moderation,⁷¹ or at worst, pose a broader threat to democratic norms.⁷²

Section 230 and the Evolution of Liability

But is Section 230 some sort of special corporate handout? A review of liability standards for carrying third-party speech shows that Section 230 is not so radically different than evolution of the common-law liability.⁷³ Section 230 means that online platforms are

not treated as publishers of third-party speech, nor do their efforts to moderate content make them liable for third-party speech. Historically though, media organizations such as newspapers were generally liable for all the speech that they published, even if it was merely republishing the speech of an outside party. As noted in the prominent legal text *Prosser and Keeton on the Law of Torts*, the traditional multiple publication rule holds that “every repetition of a defamatory statement is considered a publication,” giving rise to liability by republishers.⁷⁴

“Section 230 embodies a balance between liability and speech, recognizing the technical reality of the internet and the impossibility of perfectly monitoring every piece of user-generated content.”

But as noted by my Cato colleagues Brent Skorup and Jennifer Huddleston, the courts have been steadily diminishing this strict liability rule for republishing and distributing outside speech.⁷⁵ For example, even in the 1930s courts began to recognize that emerging technologies such as telegraph and radio presented practical difficulties of holding republishers liable for defamatory content, and so the courts began carving out new liability regimes. Over time, these radio and wire service liability regimes continued to expand to other speakers, including broader news organizations and mass media, often known as conduit liability. For example, in 1992, a plaintiff sued local CBS affiliates for running a *60 Minutes* program for supposedly defamatory content. The Ninth Circuit affirmed a district court decision in *Auvil v. CBS 60 Minutes* that holding local TV stations liable for broadcasting of *60 Minutes* content would require “the creation of full time editorial boards at local stations throughout the country which possess sufficient knowledge, legal acumen and access to experts to continually monitor incoming transmissions and exercise on-the-spot discretionary calls or face \$75 million dollar lawsuits at every turn. That is not realistic.”⁷⁶ And furthermore, the court determined that local affiliates maintained editorial control over

their broadcasts and that exercising such control at times did not mean they were liable for acting as a conduit in broadcasting third-party content.

Similarly, the Supreme Court has found that overly restrictive liability regimes would be unconstitutional because they limit the publication and access to protected speech. For example, in *Smith v. California* in 1959, the Court recognized that

the constitutional guarantees of the freedom of speech and of the press stand in the way of imposing a similar [strict liability] requirement on the bookseller. By dispensing with any requirement of knowledge of the contents of the book on the part of the seller, the ordinance tends to impose a severe limitation on the public’s access to constitutionally protected matter.⁷⁷

In 1962 the Court would continue that logic in *Manual Enterprises v. Day*, when it found:

Since publishers cannot practicably be expected to investigate each of their advertisers, and since the economic consequences of an order barring even a single issue of a periodical from the mails might entail heavy financial sacrifice, a magazine publisher might refrain from accepting advertisements from those whose own materials could conceivably be deemed objectionable by the Post Office Department. This would deprive such materials, which might otherwise be entitled to constitutional protection, of a legitimate and recognized avenue of access to the public.⁷⁸

Such decisions are important because they show that the Supreme Court recognizes that protecting free expression requires a move away from strict liability regimes and providing some allowance for distributing speech. And so Section 230, rather than being an outlier or special giveaway, actually embodies the balance between liability and speech that the courts were already developing in the common law. Section 230 flows from the same principles and recognizes the technical reality of the internet and how it is impossible for a company to fully and perfectly monitor and assess every piece of user-generated content that is posted online. The courts, including the Supreme Court, increasingly

found that the expanding reach of media required less-strict standards of liability for those new forms of communication to flourish and for the protection of legitimate speech. Section 230 merely continues that same trend. We don't fully know how courts would have handled liability on the internet, but the first two cases set perverse incentives that would have significantly reduced online speech and innovation. Given the evolution of liability before the internet, courts may very well have adapted liability law to the internet in a manner similar to Section 230. But due to the rapid rise of the internet, Section 230 created a stable liability system that allowed the further flourishing of online innovation.

The Continuing Need for Section 230

Section 230 continues to provide a clear regime that encourages innovation. Without its explicit protections, features we take for granted, such as leaving a review, engaging in debates on X, the creation and editing of information on Wikipedia, and much more would all suddenly be potential liabilities, resulting in platforms for such expression being restricted or even closed.

“The impact of removing or limiting Section 230 harms small technology companies and individuals far more than it hurts big tech companies.”

And importantly, the impact of removing or limiting Section 230 harms small technology companies and individuals far more than it hurts big tech companies. Meta, X, YouTube, and other large platforms have enough money to at least manage the compliance and legal costs in a world with a weakened Section 230. They would have to divert money away from other innovative products and services and would have to significantly suppress speech or let their platforms spin a bit out of control, but they would have the resources to at least try to manage such chaos. A smaller company has no hope of paying substantial legal costs or having the funds to both manage compliance and build new products.

So, for those concerned about Section 230 being a handout to big technology companies, repealing Section 230 would only serve to further consolidate power in the hands of existing leaders in the field. That's the reason why some large technology companies have generally been open to changes to Section 230. Yes, it will cost them some money and dynamism, but it will allow them to erect barriers that will effectively prevent any challengers from arising. One of the most important roles Section 230 plays is protecting countless start-up and smaller companies hosting and using user-generated content.

THE ROAD AHEAD FOR SECTION 230

The course of online speech both in the US and abroad shows how essential strong legal protections are for robust expression and innovation. Section 230, together with constitutional guarantees for expression, have made the US the leader in modern online innovation. But as technology continues to advance, new questions are being added to the current ones surrounding content moderation and Section 230. Artificial intelligence, for example, has created significant questions about liability that are beginning to be debated in the legal system.

So how should policymakers address these challenges? Should there be changes to Section 230 to account for concerns about harmful content online? Should Section 230 require some sense of fairness and offer a public square for all perspectives? How does Section 230 apply to AI, or does the US need a Section 230–like law for AI? What responsibility should lie with AI creators versus deployers and users?

First and foremost, policymakers should remember the powerful engine for expression and innovation that Section 230 has unleashed. Rather than pursue proposals that create regulations or incentives for either the suppression of speech or the abandonment of content moderation and curation, the core principles of Section 230 should be protected. That said, there may be proposals that could improve Section 230 without harming online innovation and speech. There is a risk, however, that opening Section 230 to reform—even if with the intent of making careful, targeted changes—may create an opportunity for significant and harmful changes. Ultimately, what might

start as a narrow reform could still carry significant risks beyond its original intent. This occurred with prior attempts, such as the Section 230 carve-out for content related to sex trafficking under FOSTA. Even what appears to be a minor change can still have significant consequences for expression or even for the existence of smaller platforms.

Section 230 Refinements

An apt place to start is to consider three changes suggested by Section 230 coauthor Chris Cox, as well as an additional related change: clarify that no viewpoint neutrality is needed under Section 230; define when platforms are considered to have developed content; limit Section 230 protections for content judged to be defamatory in a court of law; and offer clarity for unmasking anonymous speakers.⁷⁹

No viewpoint neutrality. Policymakers could clarify that platforms do not need to be unbiased or viewpoint neutral to enjoy Section 230 protections. As I have discussed in this paper, this is a frequent point of confusion despite court pronouncements of such. Perhaps the most needed reform would be to clearly state this in statute to prevent misguided attempts at contradictory laws that could limit speech.

Defining the development of content. Another proposal from Cox is to clarify when a platform creates or develops content so that it is not protected under Section 230. This recommendation makes sense if it were to be narrowly and carefully tailored. For example, this would affirm the *Roommates* decision that if a platform puts its own words into a piece of user-generated content, then it can be held liable for those words. But as we see in *Lemmon*, we must be careful that this does not create a broad product liability for hosting user-generated content.

Section 230 does not protect content that is created or developed, in whole or in part, by a platform. Congress could therefore clarify when a platform's actions make it sufficiently responsible for developing information. Minor photo filters, such as sepia or black-and-white effects, would remain protected because they do not meaningfully create content. Platform-created tools that add words, images, or substantial graphics, however, could be considered partial content creation and thus would fall outside Section 230's protections. And it is

worth reiterating that even well-intentioned reforms here risk unintended consequences that could undermine the core benefits of Section 230.

If policymakers pursue such a clarification, it should clearly distinguish between the platform creation of content and platform tools used to curate, organize, rank, or moderate it. Algorithms are tools that platforms use to carry out their expressive choices. As legal scholar Ashkhen Kazaryan notes, such curation is well protected under First Amendment jurisprudence, which safeguards not only the right to speak but also the right to curate or decline to engage in speech.⁸⁰ The Supreme Court has repeatedly affirmed this principle, such as in *Tornillo*, which upheld newspapers' discretion to host only the editorials they wished;⁸¹ *Hurley*, which defended parade organizers' right to choose participants, and emphasized that speakers do not "forfeit constitutional protection simply by combining multifarious voices";⁸² and *303 Creative*, which protect the rights of a website designer to refuse to create or curate content for views that ran contrary to her beliefs.⁸³

The Supreme Court has been clear that the First Amendment protects the rights of Americans to curate, combine, or refuse to engage in speech. A platform using an algorithm to help it scale those curation decisions to millions of users on the internet in no way weakens the fundamentally expressive nature of those decisions.

“A platform using an algorithm to help it scale curation decisions to millions of users on the internet in no way weakens the fundamentally expressive nature of those decisions.”

Not only is such algorithmic curation protected by the First Amendment, Kazaryan shows how it is also protected by Section 230. Multiple courts have agreed that algorithmic curation is protected by Section 230 because the platforms did not create the content, and so holding them liable for what they curate would be to treat them as the publishers of user-generated content—exactly what Section 230 protects against. This was the reasoning in

the Ninth Circuit’s *Dyroff v. Ultimate Software Group* and the Second Circuit’s *Force v. Facebook*, explicitly protecting algorithmic curation. And while the Supreme Court did not directly tackle Section 230 in *Gonzalez v. Google* and *Twitter v. Taamneh*, and even its decisions in *Moody v. NetChoice* and *Paxton v. NetChoice* were largely procedural, the language of these decisions was generally supportive of platforms’ rights to moderate and curate content.⁸⁴

The Third Circuit, however, departed from this approach in *Anderson v. TikTok*, a case involving a minor who died after seeing a Blackout video showing individuals choke themselves to unconsciousness. The court held that TikTok’s recommendations were its own expressive activity and therefore unprotected by Section 230.⁸⁵ This decision ignores that the source of the problem was the dangerous Blackout video produced by one of TikTok’s users. While the Court claims not to hold TikTok liable for the speech of its users, it does exactly that by arguing that any curation of speech then makes TikTok liable for the speech of its users. In so doing, the court also fails to consider the practical implications that Section 230 was created to solve. If platforms are liable for any curation of third-party content, then social media, search engines, comment sections, online marketplaces, and countless other spaces cannot safely filter or organize content in any meaningful way, resulting in a significantly worse experience for nearly all users.

Given the Supreme Court’s reluctance to clarify these issues, Congress could confirm that algorithmic tools that organize, recommend, or moderate third-party content remain protected under Section 230. Conversely, Congress can also clarify that platform-created drop-down menus, speed filters, or similar features that significantly create, alter, or change content would not be covered by Section 230. But the risk of such action may be worse than leaving such issues to the courts so that the law can develop in a way that remains protective of innovation and expression.

Defamation Rulings. A final proposal by Cox is that Section 230 protections should not apply to content that has been ruled defamatory in a state court. This judicial takedown approach attempts to avoid the problem of platforms overmoderating because of the broad risk of liability. Instead, it accepts the legal judgment of a court only on matters of defamation and creates liability for the content in question.

Such a provision, however, creates several concerns. State defamation laws vary widely.⁸⁶ While the Supreme Court imposes an “actual malice” standard for public figures, states set different standards for private individuals. They also differ on default judgments; provisions against strategic lawsuits against public participation (anti-SLAPP); statutes of limitations; available defenses; damages; retraction requirements; and whether certain statements are treated as defamatory per se.

These variations create serious adjudication and enforcement challenges. For example, a private individual in North Carolina (no anti-SLAPP law, recognizes per se defamation) could sue a user in Tennessee (strong anti-SLAPP, no per se defamation) over online speech hosted by a platform based in California (strong anti-SLAPP, recognizes per se defamation).⁸⁷ Determining which legal standard applies invites forum shopping and abuse, allowing the most restrictive states to effectively dictate online speech nationwide.

“While a judicial takedown regime is preferable to broad notice-and-takedown systems, it still poses significant risks to online expression and departs from Section 230’s core principle.”

And platforms, although not party to the underlying lawsuit, will be required to remove content or risk liability. Do platforms only need to remove the specific piece of content, or do they need to find related or identical claims? How do they handle new claims that are identical or similar to previously removed piece of content? Do they need to proactively monitor or merely reactively remove such content?

While a judicial takedown regime is preferable to broad notice-and-takedown systems, it still poses significant risks to online expression and departs from Section 230’s core principle that platforms are not “derivatively liable for the illegal acts of others.”⁸⁸

Clarity for Unmasking. An alternative to Cox’s proposal to expand platform liability for “the illegal acts of others” would be to focus on holding those actors accountable

for their own speech. This is often challenging because many platforms permit anonymous or pseudonymous participation. While pseudonyms may conceal a user’s identity from the public, platforms can typically trace users through IP addresses and other technical data, which is often sufficient to identify them. Although a minority of people use tools such as VPNs or Tor to sufficiently obscure their identity, most users can effectively be unmasked.⁸⁹

“Adopting clear standards for when anonymous speech can be unmasked would help keep the promise of Section 230.”

Any policy governing the unmasking of anonymous speech must proceed cautiously. Anonymous expression has played a vital role in American history—from colonial pamphlets to modern whistleblowing—and it enables candid discussion, exposure of wrongdoing, and participation by those who fear retaliation.⁹⁰ It protects dissidents, encourages honest reviews, and lowers barriers to civic engagement and donations.⁹¹ At the same time, anonymity can be exploited by malicious actors, underscoring the need for careful balance.

Current law on unmasking anonymous speakers varies significantly across jurisdictions and between civil and criminal contexts. One of the most speech-protective civil standards is the *Dendrite* test, developed in New Jersey. It requires plaintiffs seeking to unmask an anonymous speaker to notify the speaker, identify specific actionable statements, survive a motion to dismiss for failure to state a claim for which relief can be granted, and present prima facie evidence for each element of the claim. Even then, the court performs a balancing test weighing the speaker’s First Amendment rights against the plaintiff’s need for disclosure.

The Delaware Supreme Court adopted a different approach in *Doe v. Cahill*, which replaces *Dendrite*’s explicit balancing test with a summary judgment standard. While some courts and scholars view *Cahill* as functionally equivalent to *Dendrite*,⁹² most experts see *Dendrite* as more demanding and speech-protective.⁹³ Several jurisdictions have blended the two standards, while others apply weaker or unclear rules that make unmasking too easy.⁹⁴

These differences raise two policy questions: First, what standard should govern unmasking, and should it be set by state or federal law? There has been significant discussion in favor of *Cahill*, *Dendrite*, variations of these standards, or entirely new standards.⁹⁵ Given its strong protection for anonymous speech and its relatively wide use, *Dendrite* offers the best pro-expression option. It’s also worth noting that any reform should not require platforms to preemptively identify users, limit anonymous access through VPNs or similar tools, or undermine encryption.

The second question is how to implement such a standard. States with weak or unclear unmasking rules could clarify them through courts or legislation, allowing different approaches to develop. This state-based experimentation has led many courts to reference *Dendrite* or *Cahill*, but it has also produced a patchwork of standards that complicates enforcement and makes it harder for individuals to pursue consistent civil remedies for online speech that crosses state lines.

A single national standard would provide clarity and uniformity, but it risks being less speech-protective than *Dendrite* if Congress decides on a weaker standard. Given the borderless nature of online speech, a federal unmasking standard grounded in *Dendrite* would best protect anonymous expression while offering plaintiffs a clear, consistent framework. Such a standard could also address practical concerns, such as compensating internet service providers or platforms for compliance costs when disclosure is ultimately warranted.

Adopting clear standards for when anonymous speech can be unmasked would help keep the promise of Section 230—that the speakers of unlawful speech are to be held liable for their own speech, while society benefits from the greater online speech made possible by protecting intermediary platforms from liability.

AI and Section 230

The explosion of artificial intelligence products on the market has also raised questions about whether or not Section 230 applies to generative AI. Even among supporters of Section 230, there are a few differences of opinion. Notably, Cox and Wyden have both stated that they do not believe Section 230 applies to generative AI.⁹⁶ This does not mean that

the principles of free expression and innovation at the heart of Section 230 do not apply to AI, nor does it mean there will not be overlap between Section 230 and AI. This is because at the core of the disagreement lies the question: What is AI?

In one sense, AI is simply another class of computer tools used to solve problems and assist users. Many technologies once labeled “artificial intelligence”—from Microsoft’s Clippy to early spam filters—eventually became accepted as ordinary software. More recent generative AI tools and chatbots, however, appear different. Generative AI produces seemingly human-created content by identifying patterns across vast amounts of data, and modern chatbots use this training to engage in realistic conversations.

But critically, these AI tools only work because they have been trained on prior knowledge, that is, third-party content. The response of the AI tool or chatbot is essentially a complex summary of previous knowledge applied to answer a prompt from the user.⁹⁷ Some people therefore argue that generative AI is not meaningfully different from search engines, which organize and present third-party information. They point to cases such as *O’Kroy v. Fastcase*, where the Sixth Circuit held that Section 230 protected Google’s search-result snippets that were derived from third-party content.⁹⁸

“A stable, pro-innovation, light-touch regulatory framework would strongly benefit the AI sector.”

Others contend that generative AI participates in the development of new content, particularly as AI research aims to create increasingly autonomous, expressive systems. Proponents of this view could point to statements by Cox and Wyden as well as the apparent views of Justice Gorsuch, who argued that “Artificial intelligence generates poetry. It generates polemics today that would be content that goes beyond picking, choosing, analyzing, or digesting content. And that is not protected.”⁹⁹

So, is AI an information content provider, or is it merely part of an interactive computer service that has Section 230 protections? It is highly likely that a significant portion of the judiciary will view generative AI as taking part in creation and development of information and therefore not protected by Section 230. Even just the term “generative” is

indicative of creation and development. While it is logically consistent to view generative AI as merely the continuation of search engines and algorithms that are already protected by Section 230, it seems unlikely that this interpretation will quickly become the prevailing view in the courts, thus failing to provide sufficient protection for burgeoning AI industry.¹⁰⁰ And even if the courts do end up determining that Section 230 applies, it is likely to be a long and costly process to the AI industry before that is settled.

The key point is that generative AI may require additional legal protections to fully develop. While Section 230 already faces significant challenges, AI is an even more transformative technology, with implications for speech, medicine, industry, national security, and beyond—potentially rivaling the impact of the internet itself. A stable, pro-innovation, light-touch regulatory framework would strongly benefit the AI sector.

At the same time, AI development is highly competitive, particularly with Chinese firms building rival systems. Even policymakers skeptical of liability protections must recognize that constraining US AI companies will not slow foreign competitors or prevent bad actors from misusing AI—it will only weaken American leadership.

CONCLUSION

In this paper I have demonstrated that Section 230 has been essential to the growth of online speech and innovation. Although critics fault it for both too much and too little moderation, Section 230 has enabled platforms to operate and exercise their First Amendment rights without fear of liability for user-generated content. Efforts to weaken its core protections would create perverse incentives and harmful unintended consequences, particularly for smaller platforms.

Rather than adopting censorious or economically damaging new regulations, Section 230 should remain central to the American approach to intermediary liability that has made the United States a global leader in online expression. While narrow clarifications may improve its application, lawmakers must avoid reforms that undermine its protections. More importantly, policymakers should consider how Section 230—or a similar framework—can support and encourage AI innovation. With such protections in place, online speech enabled by new technologies can continue to sustain a free, dynamic, and prosperous society.

NOTES

1. Ash Johnson and Daniel Castro, “Fact-Checking the Critiques of Section 230: What Are the Real Problems?,” Information Technology & Innovation Foundation, February 22, 2021.
2. Cathy McMorris Rodgers and Frank Pallone Jr., “Sunset of Section 230 Would Force Big Tech’s Hand,” *Wall Street Journal*, May 12, 2024; and Dean DeChiaro, “A Real Gift to Big Tech’: Both Parties Object to Immunity Provision in USMCA,” *Roll Call*, December 17, 2019.
3. Woodrow Wilson International Center for Scholars, “Defeating Disinformation Series: Social Media Regulation Around the World,” YouTube video, February 10, 2020; and Mary Graw Leary, “The Failed Experiment of Section 230 of the Communications Decency Act: How It Facilitates Exploitation and How It Must Be Reformed,” *Villanova Law Review* 70, no. 1 (2025): 49–113.
4. “Senator Hawley Introduces Legislation to Amend Section 230 Immunity for Big Tech Companies,” Office of Senator Josh Hawley, press release, June 19, 2019; and Adam Candeub, “Bargaining for Free Speech: Common Carriage, Network Neutrality, and Section 230,” *Yale Journal of Law & Technology* 22 (2020): 391–433.
5. Jaron Lanier et al., “Sunset and Renew: Section 230 Should Protect Human Speech, Not Algorithmic Virality,” Ash Center for Democratic Governance and Innovation, Harvard Kennedy School, October 30, 2024; and Eric Goldman, “Why Section 230 Is Better than the First Amendment,” *Notre Dame Law Review Reflection* 95, no. 1 (2019): 33–46.
6. Sean Clark, “Links to the Past: a History of the World Wide Web—Timeline,” *The Guardian*, March 11, 2014; and “The Origins of the Internet,” Council on Foreign Relations, last updated January 31, 2023.
7. Jeff Kosseff, *The Twenty-Six Words That Created the Internet* (Cornell University Press, 2019).
8. *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991).
9. *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991).
10. *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991).
11. Jeff Kosseff, *The Twenty-Six Words That Created the Internet* (Cornell University Press, 2019).
12. *Stratton Oakmont Inc. v. Prodigy Services Co.*, 1995 WL 323710 (N.Y. Sup. Ct. 1995).
13. Peter H. Lewis, “Personal Computers; The CompuServe Edge: Delicate Data Balance,” *New York Times*, November 29, 1994.
14. Ellyse Dick and Ash Johnson, “What We Learned When We Broke the Internet: Insights for the Section 230 Debate,” Information Technology & Innovation Foundation, April 7, 2021.
15. Jeff Kosseff, *The Twenty-Six Words That Created the Internet* (Cornell University Press, 2019).
16. Section 230 of the Communications Act of 1934, as amended by the Telecommunications Act of 1996, 47 USC § 230. Courts have interpreted 47 USC. § 230(c)(2)(B)’s reference to “paragraph (1)” to mean § 230(c)(2)(A). See “Section 230: An Overview,” Congressional Research Service R46751, updated January 4, 2024.
17. Section 230 of the Communications Act of 1934, as amended by the Telecommunications Act of 1996, 47 USC § 230.
18. Christopher Cox, “A Section 230 Retrospective,” working paper, Center for Growth and Opportunity, November 2022.
19. Eric Goldman and Jess Miers, “Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules,” *Journal of Free Speech Law* 1, no. 1 (2021): 192–226; and “Section 230: An Overview,” Congressional Research Service R46751, updated January 4, 2024.
20. “Section 230: An Overview,” Congressional Research Service R46751, updated January 4, 2024.
21. Section 230 of the Communications Act of 1934, as amended by the Telecommunications Act of 1996, 47 USC § 230(e)(2).
22. *Reno v. ACLU*, 521 US 844 (1997).
23. *Zeran v. America Online, Inc.*, 129 F.3d 327 (4th Cir. 1997).
24. *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096 (9th Cir. 2009).
25. Evan Engstrom, “Primer: Value of Section 230,” Engine, January 31, 2019.

26. Ife Ogunleye, “Section 230: Good for Competition Online,” Chamber of Progress, *Medium*, July 5, 2022.

27. Michael Masnick, *Don’t Shoot the Message Board: How Intermediary Liability Harms Online Investment & Innovation* (Copia Institute and NetChoice, June 2019).

28. *Doe No. 1 v. Backpage.com, LLC*, No. 1:17-cv-11069 (D. Mass. 2018); and Office of Public Affairs, “Three Owners of Notorious Prostitution Website Backpage Sentenced,” US Department of Justice, August 28, 2024.

29. *Gonzalez v. Google LLC.*, 598 US 617 (2023); and *Twitter, Inc. v. Taamneh*, 598 US 471 (2023).

30. See also cases such as the Rohingya, a Muslim minority group in Myanmar that faced a genocide starting in 2016 at the hands of the Myanmar military and other ethnic and religious groups. The UN estimates that at least 25,000 Rohingya were killed and hundreds of thousands displaced. Facebook was a prominent source of news and connection within Myanmar and became a place where anti-Rohingya views and conspiracies frequently spread. Various groups have blamed Meta for its role in inflaming and contributing to the genocide against the Rohingya. “United States: Rohingya Survivor Asks US Regulator to Investigate Meta’s Potential Role in Myanmar Atrocities,” Amnesty International, January 23, 2025; and *Doe v. Meta Platforms, Inc.*, No. 4:22-cv-00051-ygr, (N.D. Cal. 2024).

31. Laura Moraff, “How Online Censorship Harms Sex Workers and LGBTQ Communities,” ACLU, June 27, 2022.

32. US Government Accountability Office, “Sex Trafficking: Online Platforms and Federal Prosecutions,” GAO-21-385, June 2021; and *G. G. v. Salesforce.com, Inc.*, No. 22-2621 (7th Cir 2023).

33. Fentanyl Trafficking Prevention Act, S. 2264, 118th Cong. (2023); and Deplatform Drug Dealers Act, H.R. 1599, 118th Cong. (2023).

34. Accountability for Online Firearms Marketplaces Act of 2021, S. 2725, 117th Cong. (2021).

35. See Something, Say Something Online Act of 2021, S. 27, 117th Cong. (2021).

36. See Something, Say Something Online Act of 2021, S. 27, 117th Cong. (2021).

37. SAFE TECH Act, S. 560, 118th Cong. (2023).

38. Justice Against Malicious Algorithms Act of 2021, H.R.

5596, 117th Cong. (2021); and Protecting Americans from Dangerous Algorithms Act, H.R. 2154, 117th Cong. (2021).

39. Federal Big Tech Tort Act, H.R. 5449, 117th Cong. (2021).

40. Danielle Keats Citron and Mary Anne Franks, “The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform,” Boston University School of Law, *Public Law Research Paper* no. 20-8 (February 2020); and Danielle Keats Citron and Benjamin Wittes, “The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity,” University of Maryland Francis King Carey School of Law, *Legal Studies Research Paper* no. 2017-22 (July 27, 2017).

41. A Bill to Repeal Section 230 of the Communications Act of 1934, S. 5020, 116th Cong. (2020); Sen. Josh Hawley (@HawleyMO), “Time to FULLY repeal Section 230. Open the courtroom doors. Allow the people to SUE the AI companies that take away their rights,” X post, September 4, 2025, 11:56 a.m.; and Amy Klobuchar, “What I Didn’t Say About Sydney Sweeney,” *New York Times*, August 20, 2025.

42. Nico Perrino (@nicoperrino), “It is a bedrock principle of the First Amendment that the government may not prohibit the expression of an idea merely because it is offensive. ... ,” X post, September 17, 2025, 9:04 a.m.

43. Jess Miers, “Your Problem Is Not with Section 230, But the 1st Amendment,” *Techdirt*, November 2, 2020.

44. Section 230 of the Communications Act of 1934, as amended by the Telecommunications Act of 1996, 47 USC § 230(f)(3).

45. *Fair Housing Council of San Fernando Valley v. Roommates.com LLC*, 521 F.3d 1157, 1168 (9th Cir. 2008) (en banc).

46. *Herrick v. Grindr LLC*, 18-396 (2d Cir. 2019); Will Duffield, “Circumventing Section 230: Product Liability Lawsuits Threaten Internet Speech,” Cato Institute Policy Analysis no. 906, January 26, 2021; and Carrie Goldberg, “Winning Through Losing,” *ABA*, December 18, 2020.

47. *Lemmon v. Snap, Inc.*, 995 F.3d 1085 (9th Cir. 2021).

48. Eric Goldman, “The Ninth Circuit’s Confusing Ruling Over Snapchats Speed Filter—Lemmon v. Snap,” *Technology & Marketing Law Blog* (blog), May 12, 2021.

49. Section 230 of the Communications Act of 1934, as amended by the Telecommunications Act of 1996, 47 USC § 230(f)(3).

50. While many on the right are quick to see many instances

of content moderation as evidence of bias, it's also true that in plenty of cases platforms created policies that favored progressive views. For example, Meta's fact-checking regime would be used to suppress the Hunter Biden laptop story published by the *New York Post* in the final days of the 2020 election. Or consider Twitter's 2018 adoption of policies that prohibit misgendering or deadnaming transgender individuals, which accepts a view of sexuality and gender that many on the right or those with traditional religious beliefs would reject. But even these examples show how ideological choices made by social media companies brought change. The Hunter Biden laptop story, together with other examples where Meta's fact-checking program or misinformation policies got things wrong, and increasing user frustration with the fact-checking program led Meta to shutter the program in 2025 and roll back many of its misinformation policies. The Twitter rule resulted in high-profile criticisms, notably by satirical organizations such as the *Babylon Bee*, which lost its Twitter account for calling Health and Human Services Assistant Secretary Rachel Levine a man. But the banning of the *Babylon Bee* appears to be the final straw that ultimately led Elon Musk to buy Twitter and reshape the platform. The point is that bias does exist, but companies can and must respond to demands of users or else lose out to revolt and competition. See also Cato polling showing that conservatives are more likely to report having their content penalized by social media companies: David Kemp and Emily Ekins, "Poll: 75% Don't Trust Social Media to Make Fair Content Moderation Decisions, 60% Want More Control over Posts They See," Cato Institute Survey Report, December 15, 2021.

51. "Governor Abbott Signs Law Protecting Texans from Wrongful Social Media Censorship," Office of the Texas Governor, September 9, 2021.

52. CBS News, "Sen. Ted Cruz Grills Mark Zuckerberg on Political Bias," YouTube video, April 10, 2018.

53. Donald Trump (@realDonaldTrump), "Republicans feel that Social Media Platforms totally silence conservatives voices. We will strongly regulate, or close them down, before we can ever allow this to happen. We saw what they attempted to do, and failed, in 2016. We can't let a more sophisticated version of that....," X post, May 27, 2020, 7:11 a.m.

54. Protect Speech Act, H.R. 8517, 116th Cong. (2020); Online Freedom and Viewpoint Diversity Act, S. 4534, 116th Cong. (2020); Preserving Political Speech Online Act, S. 2338, 117th Cong. (2021); and DISCOURSE Act, S. 921, 118th Cong. (2023).

55. Stop the Censorship Act, H.R. 908, 119th Cong. (2025).

56. 21st Century FREE Speech Act, S. 1384, 117th Cong. (2021).

57. A Bill to Repeal Section 230 of the Communications Act of 1934, S. 2972, 117th Cong. (2021); and AOC Act, H.R. 874, 117th Cong. (2021).

58. Florida, An Act Relating to Social Media Platforms, S.B. 7072, Regular Session (2021).

59. Texas, An Act Relating to Censorship of or Certain other Interference with Digital Expression, Including Expression on Social Media Platforms or Through Electronic Mail Messages, H.B. 20, 87th Leg., 2nd Called Sess. (2021).

60. Jennifer Huddleston, "Why the NetChoice Ruling Is a Win for Free Speech," Cato Institute, July 1, 2024.

61. *Moody v. NetChoice, LLC*, 603 S. Ct. __ (2024).

62. Brief for Petitioners, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023); Brief of amici curiae Bluesky et al. in support of NetChoice and CCIA, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023); Brief of amici curiae Chamber of Progress et al. in support of NetChoice, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023); Brief of amicus curiae Discord Inc. in support of NetChoice, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023); Brief of amici curiae Cato Institute et al. in Support of NetChoice, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023); and Brief of amici curiae Marketplace Industry Association et al. in Support of NetChoice, *NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court 2023).

63. Brief of amici curiae Trust & Safety Foundation in Support of Petitioners, *NetChoice, LLC v. Paxton*, Nos. 22-277 & 22-555 (US Supreme Court 2023).

64. Thomas M. Johnson Jr., "The FCC's Authority to Interpret Section 230 of the Communications Act," Federal Communications Commission, October 21, 2020.

65. David Inserra and Brent Skorup, "Comments in Response to FCC 'Delete, Delete, Delete' Initiative," Cato Institute, April 11, 2025.

66. *Loper Bright Enterprises v. Raimondo*, 603 US __ (2024); Lawrence J. Spiwak, "Sauce for the Goose: The FCC Lacks Authority to Interpret Section 230 Post-Loper Bright," Federalist Society, November 21, 2024; and Lawrence J. Spiwak, "The FCC Still Can't Interpret Section 230," Federalist Society, May 1, 2025.

67. “Schakowsky, Castor Reintroduce Legislation to Hold Online Platforms Accountable,” Office of Representative Jan Schakowsky, press release, April 15, 2025; State of the Net, “The Frontiers for Free Speech in the Trump Era: The Content Moderator’s Dilemma,” Youtube video, February 11, 2025; and Brief for amicus curiae the Digital Progress Institute in Support of Respondents, *NetChoice, LLC v. Moody & NetChoice, LLC v. Paxton*, Nos. 22-277 and 22-555 (US Supreme Court).

68. “Terms of Service,” Meta, January 1, 2025.

69. See *Prager University v. Google LLC*, 951 F.3d 991 (9th Cir. 2020), *Calise v. Meta Platforms, Inc.* 4:21-cv-06186-JSW (N.D. Cal. Sept. 22, 2025); *Glazer v. Meta Platforms, Inc.*, 2025 WL 2958810 (D. Md. Oct. 17, 2025) and a discussion of the differences in court rulings on this by Eric Goldman, Section 230 Applies to Scammy Ads—Galzer v. Facebook, *Technology & Marketing Law Blog* (blog), October 23, 2025.

70. Cathy McMorris Rodgers and Frank Pallone Jr., “Sunset of Section 230 Would Force Big Tech’s Hand,” *Wall Street Journal*, May 12, 2024; and Mike Masnick, “While Democracy Burns, Democrats Prioritize Demolishing Section 230,” *Techdirt*, February 21, 2025.

71. Robert Winterton, “Calling Section 230 a ‘Big Tech Handout’ Is a Lie. Here’s Why,” *NetChoice*, July 29, 2019.

72. Donald Trump (@realDonaldTrump), “Section 230, which is a liability shielding gift from the U.S. to ‘Big Tech’ (the only companies in America that have it - corporate welfare!), is a serious threat to our National Security & Election Integrity. Our Country can never be safe & secure if we allow it to stand.....,” X post, December 1, 2020, 9:45 p.m.

73. Brent Skorup and Jennifer Huddleston, “The Erosion of Publisher Liability in American Law, Section 230, and the Future of Online Curation,” working paper, Mercatus Center, July 11, 2019.

74. *Zeran v. America Online*, 129 F.2d 327, 332 (4th Cir. 1997) citing W. Page Keeton, ed., *Prosser and Keeton on the Law of Torts*, 5th ed. (West Publishing, 1984).

75. Brent Skorup and Jennifer Huddleston, “The Erosion of Publisher Liability in American Law, Section 230, and the Future of Online Curation,” working paper, Mercatus Center, July 11, 2019.

76. *Auvil v. CBS 60 Minutes*, 800 F. Supp. 928 (E.D. Wash. 1992) at 931; and *Auvil v. CBS 60 Minutes*, 67 F.3d 816 (9th Cir. 1995).

77. *Smith v. California*, 361 US 147 (1959).

78. *Manual Enterprises, Inc. v. Day*, 370 US 478 (1962).

79. Christopher Cox, “A Section 230 Retrospective,” working paper, Center for Growth and Opportunity, November 2022.

80. Ashkhen Kazaryan, “I Built This Algo Brick by Brick: Why You Should Care About Algorithms, Free Speech, and Section 230,” *The Future of Free Speech*, September 10, 2025.

81. *Miami Herald Publ’g Co. v. Tornillo*, 418 US 241 (1974).

82. *Hurley et al. v. Irish-American Gay, Lesbian and Bisexual Group of Boston, Inc., et al.*, 515 US 557, 574 (1995).

83. “All manner of speech . . . qualify for the First Amendment’s protections; no less can hold true when it comes to speech like Ms. Smith’s conveyed over the Internet.” *303 Creative LLC v. Elenis*, 600 US ___ (2023).

84. Ashkhen Kazaryan, “I Built This Algo Brick by Brick: Why You Should Care About Algorithms, Free Speech, and Section 230,” *The Future of Free Speech*, September 10, 2025.

85. *Anderson v. TikTok, Inc.*, No. 22-3061 (3rd Cir. 2024).

86. Gibson, Dunn & Crutcher LLP, “Q&A: Defamation Claims in USA,” *Lexology*, June 30, 2023; David Keating, Helen Knowles-Gardner, and Dan Greenberg, “Anti-SLAPP Statutes: 2025 Report Card,” *Institute for Free Speech*, August 25, 2025; Danielle Gilmore, Esq., “What Is Defamation Per Se?,” *FindLaw*, August 17, 2023; and David Goguen, “Defamation, Libel, and Slander Statute of Limitations,” *Nolo*, updated March 26, 2024.

87. David Keating, Helen Knowles-Gardner, and Dan Greenberg, “Anti-SLAPP Statutes: 2025 Report Card,” *Institute for Free Speech*, August 25, 2025.

88. Christopher Cox, “A Section 230 Retrospective,” working paper, Center for Growth and Opportunity, November 2022.

89. Leeza Arbatman and John Vilasenor, “Anonymous Expression and ‘Unmasking’ in Civil and Criminal Proceedings,” *Minnesota Journal of Law, Science & Technology* 23, no. 1 (2022): 77–130.

90. I would be remiss if I did not specifically mention *Cato’s Letters*, written pseudonymously by John Trenchard and Thomas Gordon. Most famously they wrote in *Cato’s Letter* no. 15 that “Freedom of speech is the great bulwark of liberty; they prosper and die together: And it is the terror of traitors

and oppressors, and a barrier against them. It produces excellent writers, and encourages men of fine genius.”

91. “Anonymity,” Electronic Frontier Foundation.

92. *Doe v. Cahill*, 884 A.2d 451 (Del. Super. 2005).

93. Matthew Mazzotta, “Balancing Act: Finding Consensus on Standards for Unmasking Anonymous Internet Speakers,” *Boston College Law Review* 51, no. 3 (2010): 833–870; Jessica L. Chilson, “Unmasking John Doe: Setting a Standard for Discovery in Anonymous Internet Defamation Cases,” *Virginia Law Review* 95, no. 2 (2009): 389–433.

94. Leeza Arbatman and John Vilasenor, “Anonymous Expression and ‘Unmasking’ in Civil and Criminal Proceedings,” *Minnesota Journal of Law, Science & Technology* 23, no. 1 (2022): 77–130.; and author research into state court unmasking standards.

95. Ben Sperry, “Right to Anonymous Speech, Part 2: A Law & Economics Approach,” Truth on the Market, September 6, 2023; Leeza Arbatman and John Vilasenor, “Anonymous Expression and ‘Unmasking’ in Civil and Criminal Proceedings,” *Minnesota Journal of Law, Science*

& Technology 23, no. 1 (2022): 77-130; and Shelly L. Skeen, “Unmasking Anonymous Internet Speakers: Balancing the Right of Free Speech Against the Right to Seek Redress for Injury for Wrongful Online Speech,” *On the Cusp* 1, no. 3 (2019).

96. Cristiano Lima-Strong, “AI Chatbots Won’t Enjoy Tech’s Legal Shield, Section 230 Authors Say,” *Washington Post*, March 17, 2023.

97. Jess Miers, “Yes, Section 230 Should Protect ChatGPT and Other Generative AI Tools,” Techdirt, March 17, 2023.

98. Jess Miers, “Yes, Section 230 Should Protect ChatGPT and Other Generative AI Tools,” Techdirt, March 17, 2023.

99. Matt Perault, “Section 230 Won’t Protect ChatGPT (or Other AI-Generated Content),” *Journal of Free Speech Law* 3, no. 2 (2023); and Cristiano Lima-Strong, “AI Chatbots Won’t Enjoy Tech’s Legal Shield, Section 230 Authors Say,” *Washington Post*, March 17, 2023.

100. Natallie Rocha, “Google and Character.AI to Settle Lawsuit Over Teenager’s Death,” *New York Times*, Jan 7, 2026.

RELATED PUBLICATIONS FROM THE CATO INSTITUTE

The Misleading Panic over Misinformation: And Why Government Solutions Won’t Work by David Inserra, Policy Analysis no. 999 (June 26, 2025)

Content Creators, Entrepreneurial Users, and the Impact of Tech Policy by Jennifer Huddleston, Policy Analysis no. 995 (April 29, 2025)

A Guide to Content Moderation for Policymakers by David Inserra, Policy Analysis no. 974 (May 21, 2024)

Competition and Content Moderation: How Section 230 Enables Increased Tech Marketplace Entry by Jennifer Huddleston, Policy Analysis no. 922 (January 31, 2022)

Circumventing Section 230: Product Liability Lawsuits Threaten Internet Speech by Will Duffield, Policy Analysis no. 906 (January 26, 2021)

RECENT STUDIES IN THE CATO INSTITUTE POLICY ANALYSIS SERIES

- 1012. Politically Motivated Killers: 51 Years of Terrorist Murders on US Soil, 1975–2025** by Alex Nowrasteh (February 19, 2026)
- 1011. COVID-19 Slowed but Couldn't Stop the Fall in Global Inequality: Updating the Inequality of Human Progress Index** by Chelsea Follett and Vincent Geloso (February 3, 2026)
- 1010. Ending Community Development Subsidies** by Chris Edwards (January 15, 2026)
- 1009. Understanding Debanking: Evaluating Governmental, Operational, Political, and Religious Financial Account Closures** by Nicholas Anthony (January 8, 2026)
- 1008. Better Long-Term Care for Billions Less** by Stephen A. Moses (November 13, 2025)
- 1007. Steeled for Protectionism** by Clark Packard and Alfredo Carrillo Obregon (October 30, 2025)
- 1006. Their Own Two Feet: How Europeans Can Lead on Conventional Deterrence in Europe** by Benjamin David Giltner (October 23, 2025)
- 1005. Licensing Requirements Would Block Care and Justice for Sexual Assault Victims** by Jeffrey A. Singer, Akiva Malamet, Spencer Pratt, and Michael F. Cannon (October 1, 2025)
- 1004. The \$15 Trillion Emergency Spending Loophole Federal Emergency Spending: 1991–2025** by Dominik Lett (September 16, 2025)
- 1003. Immigrants Cut Victimization Rates, Boost Crime Reporting** by David J. Bier and Julián Salazar (August 26, 2025)
- 1002. Reforming Federal Laws on Private Sector Labor Unions** by Charles W. Baird (July 29, 2025)
- 1001. Unnecessary Relics: The Surgeon General and the Public Health Service Commissioned Corps** by Jeffrey A. Singer, Akiva Malamet, Bautista Vivanco, and Michael F. Cannon (July 22, 2025)
- 1000. Aimless Rivalry: The Futility of US–China Competition in the Middle East** by Jon Hoffman (July 10, 2025)
- 999. The Misleading Panic over Misinformation: And Why Government Solutions Won't Work** by David Inserra (June 24, 2025)

CITATION

Inserra, David. "The Future of Online Expression and Innovation Depends on Robust Section 230 Protections," Policy Analysis no. 1013, Cato Institute, Washington, DC, February 26, 2026.



The views expressed in this paper are those of the author(s) and should not be attributed to the Cato Institute, its directors, its Partners, or any other person or organization. Nothing in this paper should be construed as an attempt to aid or hinder the passage of any bill before Congress. Copyright © 2026 Cato Institute. This work by the Cato Institute is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.